



PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Mark WASSON et al.	:	Confirmation No.: 4117
Application No. 10/716,202	:	TC/Art Unit: 2626
Filed: November 19, 2003	:	Examiner: Paras Shah
For: SYSTEM AND METHOD FOR	:	Atty Docket: P68795US0
EXTRACTING INFORMATION FROM	:	
TEXT USING TEXT ANNOTATION	:	
AND FACT EXTRACTION (<i>as</i>	:	
<i>amended</i>)	:	

DECLARATION UNDER 37 CFR § 1.132

Mail Stop AMENDMENT
Commissioner for Patents
PO Box 1450
Alexandria, VA 22313-1450

Sir:

I, MARK WASSON, declare that:

1. I am one of the named inventors of this application. My Curriculum Vitae is attached hereto.
2. One of the features of the invention that is the subject of this application is that annotated text is represented in a single view of a document expressed as inline XML. This feature, although not explicitly described in the application as filed, is necessarily present in the invention as described in the application as filed, and that it would be so recognized by persons of ordinary skill.
3. In particular, a person of ordinary skill in the art would recognize this feature to be necessarily present in the invention as described in the application as originally filed, because

all of the descriptions or examples in the application show FEX annotations captured in single representations, including the representation used in the description of the uncross algorithms. None of the descriptions or examples in the application shows FEX annotations captured in two or more document views. Specific instances of such descriptions and examples showing FEX annotations captured in single representations (with reference to the paragraph numbers and Figures in the application as filed) and my accompanying comments are as follows:

Paragraph number(s)	Comment(s)
85-89	These paragraphs describe how multiple annotators can introduce conflicts. The specific XML examples in paragraphs 87 and 88 are within single representations of the text. Furthermore, the scope of XML "well-formedness" is only within a single representation.
74-77	These go from a single tree representation to a single XML representation.
124-125	Examples with overlapping annotations before and after uncrossing. The annotations are captured in a single document view.
64-65	The token "Mary" is annotated with a number of annotations, and it has a location in a parse tree. All of this is captured in a single graphical representation, the tree shown in Figure 2.
59	The description is couched in terms of the singular use of "tree" and "representation." (Text: "One typical characteristic of parser-based text annotations is that the annotations are usually represented by a tree or some other hierarchical representation.")
99	This paragraph explains how the invention accomplishes annotation. We (the inventors) discuss this as "an XML representation [singular], and describe how token attributes, constituent attributes and shifts are handled within that representation. (Text: "In an XML representation, token attributes tend to be represented as XML attributes on base tokens, and constituent attributes and links tend to be represented as XML elements. Shifts tend to be represented as XPath expressions that utilize token attributes, constituent attributes, and links.")
105	This paragraph includes another reference to Figure 2

Paragraph number(s)	Comment(s)
116	This paragraph describes our (the inventors') reasons for choosing to use XML, including how it can capture tree structure and constituent-specific attributes. Twice we refer to "an XML-based representation [singular] of the document."
119	Again we refer to "representation" [singular] of a tree structure [also singular]. (Text: "a token may have a few or a few dozen attributes assigned to it, directly or through its parents in the tree structure representation ("representation" referring here to the fact that what is stored on the computer is a representation of a tree structure).")
177	Our pattern recognition language applies to a "text representation" [singular]. (Text: "The RuBIE pattern recognition language can exploit any attributes with which a text representation has been annotated.")
203	We focus the application of a pattern recognition rule on an XML-based annotation representation [singular]. (Text: "No sort of taxonomical inheritance is assumed; otherwise a pattern recognition rule would have to draw information from sources in addition to the XML-based annotation representation.")

4. In view of my foregoing comments, it is my opinion that a person of ordinary skill in the art would recognize from reading the application as filed that the representation of annotated text in a single view or representation of a document expressed as inline XML, is an inherent feature of and necessarily present in the invention.

5. I understand that willful false statements and the like are punishable by fine or imprisonment, or both (under 18 U.S.C. § 1001) and may jeopardize the validity of the application or any patent issuing thereon.

6. All statements made of my own knowledge are true and that all statements made on information and belief are believed to be true.

Date: April 24, 2009

Signed: Mark Wasson
Mark Wasson



Curriculum Vitae for Mark David Wasson

Work History

LexisNexis (originally Mead Data Central), 1986-Present
Current Title: Senior Architect (since 2000)

The University of Iowa Department of Linguistics (1984-1986)
Teaching Assistant

The University of Iowa Department of Preventive Medicine (1983-1986)
Programmer

Past and Present R&D Areas at LexisNexis

- Named entity extraction, classification and linking
- Document classification and indexing
- Fact extraction
- Content integration and linking
- Knowledge discovery in text-based metadata
- Trend analysis
- Sentence boundary detection
- Automatic summarization
- Duplicate document detection

Other LexisNexis-related Work Activities

- Collaborative R&D with 3rd parties
- 3rd party site visits, scouting reports and technical evaluations
- Monitoring and reporting on related research, technical and trade literature
- Competitive intelligence analysis
- Technical due diligence support for Reed Elsevier Ventures
- Technical due diligence support for acquisition candidates

Education

The University of Iowa

- Master of Arts, Linguistics (1986)
- Bachelor of Science, Computer Science (1984)
- Bachelor of Arts, Linguistics (1984)

Refereed Papers and Presentations

Baldwin, B., Doran, C., Reynar, J., Srinivas, B., Niv, M., & Wasson, M. (1997). *EAGLE: An Extensible Architecture for General Linguistic Engineering*. RIAO 97 Conference Proceedings.

Wasson, M. (1998). *Using Leading Text for News Summaries: Evaluation Results and Implications for Commercial Summarization Applications*. COLING-ACL '98 Conference Proceedings.

Shewhart, M., & Wasson, M. (1999). *Monitoring a Newsfeed for Hot Topics*. Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Wasson, M. (2000). *Using Document Metadata for Information Retrieval and Knowledge Discovery*. Poster Presentation. Proceedings of the Society of Competitive Intelligence Professionals 15th Annual International Conference and Exhibit.

Wasson, M. (2000). *Large-scale Controlled Vocabulary Indexing for Named Entities*. Proceedings of the Language Technology Joint Conference: ANLP-NAACL 2000.

Wasson, M., & Daley, R. (2000). *Searching and Access to Full Content on the Web, or, We've Got Documents and Publications, Now What?* The Serials Librarian, Volume 38, Numbers 1/2.

Wasson, M., & Daley, R. (2000). *Searching and Access to Full Content on the Web, or, We've Got Documents and Publications, Now What?* In P. M. Fiander, J. C. Harmon, & J. D. Makepeace (eds.), From Carnegie to Internet2, Forging the Serials Future, Haworth Information Press: Binghamton, NY.

Wasson, M. (2001). *Classification Technology at LexisNexis*. SIGIR 2001 Workshop on Operational Text Classification.

Wasson, M. (2001). *Examining Categorization Technologies in the Context of Product Requirements*. Panel Presentation. 64th American Society for Information Science and Technology Annual Meeting.

Wasson, M. (2002). *Using Summaries in Document Retrieval*. Proceedings of the 2002 Workshop on Automatic Summarization.

Wasson, M. (2004). *Connecting the Docs: Integrating Information from Multiple Documents*. Presentation. 2004 American Society for Information Science and Technology Pacific Northwest Chapter Annual Meeting.

Other Papers and Presentations

Zhou, J., & Wasson, M. (1993). *Processing Large Text Corpora: Shallow vs. Deep*. Invited Talk. Wright State University Computer Science Department Colloquium.

Wasson, M. (1993). *Challenges for Large Corpora Processing in Industry*. Invited Talk. First Workshop on Very Large Corpora: Academic and Industrial Perspectives.

Wasson, M. (1993). *Large Corpora Processing in Industry*. Presentation at the University of Pennsylvania, November 16, 1993.

Baldwin, B., Reynar, J., Srinivas, B., Doran, C., Niv, M., & Wasson, M. (1996). *NLP Spade: An Extensible Tool for Natural Language Engineering*. Manuscript.

Zhou, J., Meier, E., Wassum, R., Zhang, P., Wasson, M., Parhizgar, A., & Pliske, D. (1998). *LEXIS-NEXIS Statistical Summarization Research Description*. Presentation. Summarization Automation Conference SUMMAC-1, TIPSTER Text Phase III 18-month Workshop Proceedings.

Wasson, M. (1999). *University R&D and Industry R&D and Commercialization*. Panel Presentation. National Science Foundation Information and Data Management Workshop '99: Research Agenda for the 21st Century.

Wasson, M., & Daley, R. (1999). *Breaking the Database Barrier: We've Got Documents and Publications, Now What?* Presentation. North American Serials Interest Group NASIG 14th Annual Conference.

Wasson, M. (1999). *Preparing Texts for Knowledge Management and Knowledge Discovery*. Presentation. IQPC Corporate Portals Conference.

Wasson, M. (2002). *Data Mining and Text-based Information*. Presentation. Southern Ohio Chapter of the American Society for Information Science & Technology.

Wasson, M. (2003). *Text Processing and Retrieval at LexisNexis*. Presentation at Microsoft Research, April 4, 2003.

Wasson, M. (2003). *NLP R&D and Commercial Deployment*. Tutorial. 2003 Human Language Technology-North American Chapter of the Association for Computational Linguistics Joint Conference.

Wasson, M. (2003). *Classify This! Classification Today and Tomorrow*. Invited Talk. 3rd Workshop on Operational Text Classification.

Conference and Workshop Panel Presentations

Croft, B., Maier, D., Wasson, M., & Salzberg, B. (1999). *University R&D and Industry R&D and Commercialization*. Panel Discussion and Presentations. National Science Foundation Information and Data Management Workshop '99: Research Agenda for the 21st Century.

Grove, A., Thompson, P., Wasson, M., Schuetze, H., & Dornfield, E. (2001). *Categorization in Industry: Case Studies*. Panel Discussion. 64th American Society for Information Science and Technology Annual Meeting.

Conference and Workshop Organization Activities

4th Workshop on Very Large Corpora (1996). Program committee.

Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora (2000). Program committee.

2001 Workshop on Operational Text Classification (2001). Program committee.

2003 Workshop on Operational Text Classification (2003). Organizing committee.

2005 Conference on Using Metadata to Manage Unstructured Text (2005). Review panel.

32nd Annual ACM SIGIR Conference (2009). Program committee.

Other Professional Activities

Advisor, University of Pennsylvania 6th Message Understanding Conference (MUC-6) team (1995). See Baldwin, B., Reynar, J., Collins, M., Eisner, J., Ratnaparkhi, A., Rosenzweig, J., Sarkar, A., & Srinivas, B. (1995). *University of Pennsylvania: Description of the University of Pennsylvania System Used for MUC-6*. Proceedings of the Sixth Message Understanding Conference (MUC-6).

Member, LexisNexis Summarization Automation Conference (SUMMAC) team (1998). See Zhou, J., Meier, E., Wassum, R., Zhang, P., Wasson, M., Parhizgar, A., & Pliske, D. (1998). *LEXIS-NEXIS Statistical Summarization Research Description*. Presentation. Summarization Automation Conference SUMMAC-1, TIPSTER Text Phase III 18-month Workshop Proceedings.

Member, IDM 99 Languages, Data Models and Application Models Workshop Discussion Group, National Science Foundation Information and Data Management Workshop '99: Research Agenda for the 21st Century (1999). See Urban, S., & Belkin, N. (co-chairs) (1999). *Languages, Data Models and Application Models Workshop Discussion Group Report*. <http://www.cs.ucla.edu/csd/IDM99/FinalReports/#3>

Industry Representative, NAACL HLT Advisory Board (2007-present).

Professional Affiliations

Association for Computational Linguistics

Association for Computing Machinery

- ACM Special Interest Group on Artificial Intelligence
- ACM Special Interest Group on e-Commerce
- ACM Special Interest Group on Information Retrieval
- ACM Special Interest Group on Knowledge Discovery in Data and Data Mining

Association for the Advancement of Artificial Intelligence

International Association for Artificial Intelligence and Law

Since 1987 I have attended more than 70 conferences, workshops and symposia, most of which were associated with one or more of the above organizations.